DOCUMENT RESUME

ED 336 945 FL 019 380

AUTHOR Mason, Victor W.; And Others
TITLE A New EFL Proficiency Test for Kuwait University.
PUB DATE 91
NOTE 32p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Educational Improvement; *English (Second Language); Foreign Countries; Higher Education; Item Banks; *Language Proficiency; *Language Tests; Second Language Instruction; *Student Placement; *Test Construction; Test Format; Test Items; Test Validity
IDENTIFIERS Kuwait; *Kuwait University

ABSTRACT

Kuwait University initially developed a new English placement test for students to help upgrade English language instruction throughout the country's educational system, through diagnosis of relative English proficiency levels of all secondary school graduates, determination of suitable admissions criteria for all higher education institutions, and appropriate placement of new college students. At first the target population was the approximately 22,000 secondary school students graduating annually, and test content was based on the text series used in the schools. Feasibility considerations reduced the target population to 3,000-3,500 entering Kuwait University students. The test included three parts: grammatical structure, vocabulary, and reading comprehension. A four-option multiple-choice format was selected for all three, based on careful consideration of its limitations and potential. Language center staff wrote items according to test-content specifications and item-writing guidelines. A bank of almost 2,000 new items was pre-tested and refined, and the first full-length form was piloted at the university in 1987, then administered to secondary school students. Results were quite varied, and revealed patterns in proficiency and college admission. Subsequent test development has allowed further refinement of items and content. All participants in the development process benefitted. Statistical results are appended. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED336945

A New EFL Proficiency Test for Kuwait University

by Victor W. Mason

Ali S. Hajjaj

Balkees Al-Najjar

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Mason, Victor W
& others

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL019380

BEST COPY AVAILABLE

Abstract

The New English Placement Test project of Kuwait University Language Center began as an effort to help upgrade English-language education throughout the Kuwait educational system: from middle school through university. Although the original ambitious scope of the project later had to be scaled back, an enormous amount of excellent test-development work was accomplished by the KULC administration and teaching staff, with the valued aid of a team of psychometric experts.* Achievements included creation of a huge bank of test items and reading passages, as well as mountains of statistical data on secondary-school and university student populations, as well as on item, part and test-form effectiveness. Data gleaned and lessons in responsible testing practice learned can be of value to foreign-language teaching programs everywhere.

*The authors wish to take this opportunity to express their appreciation of the assistance lent to the project in 1987-90 by the three psychometric experts who gave so generously of their time to the staff of Kuwait University Language Center: Dr. Ragaa Abou-Allam, of the College of Education, Kuwait University; Dr. Robert Ziomek, chief of evaluation of the Cedar Rapids, Iowa (U.S.A.), public school system; and Dr. Salma Khammash, of the University of Michigan.

A New EFL Proficiency Test for Kuwait University

I. Introductory Comments

The authors would like to thank Prof. Maria Piotrowska and the organizing committee of the Language Symposium at the University of Qatar for inviting us to participate in this conference. We are disappointed not to be able to attend personally, as originally anticipated, but are very grateful that a presenter has been found to deliver the report in our stead.

We are, of course, saddened that recent political developments in the Gulf region have made our attendance at the symposium impossible. Moreover, the test development project reported here is, at present, moot, until Kuwait University, like Kuwait as a society and nation, can recover from the year's traumatic events.

Nevertheless, there is a kind of solemn consolation derived from memories of the civilized institutions, like education, that Kuwait, as a modern peace-abiding state, was devoting its resources to, for the betterment of its people. Those of us privileged over the years to have

---

Dr. Balkees Al-Najjar launched the NEPT project during her tenure as KULC director from 1985 until 1989. Dr. Ali Hajjaj headed the NEPT project in his capacity as KULC assistant director between 1987 and 1989. Mr. Victor W. Mason served as acting head of the KULC Tests & Measurement Office between 1980 and 1990. Any errors of fact or interpretation are the sole responsibility of Mr. Mason.

played a significant role in Kuwait's development in projects like that reported here are grateful to have had the opportunity to make our own personal contributions.

Finally, it is because this KULC testing project was expected eventually to affect cooperative efforts in higher education among the GCC states that we feel that this paper remains relevant to the interests of participants in this symposium. May Kuwait University Language Center be remembered for the dedication of its staff, Kuwaiti and non-Kuwaiti, to the principles of sound foreign-language teaching, testing and research. With luck, insha'alah, the academic programs of Kuwait University will continue, once peace has been restored in this region.

II. Origins of the NEPT Project

A. New Kuwait University Academic Standards

KULC embarked in late 1986 on its massive New English Placement Test (NEPT) project as the university's new administration undertook markedly to upgrade academic standards, in part by reducing the annual intake of students, as well as by introducing stiffer entry requirements for first-year students in the nine undergraduate colleges. The expectation was that batteries of new standardized admissions tests would be developed by the university in key subject areas like Arabic, mathematics, English as a foreign language (EFL),

and perhaps science. Secondary-school graduates believed to show academic promise but scoring below predetermined cut-off points on the tests would be required to take at least a year of remedial studies in the subject areas in which they were considered deficient.

B. NEPT Project Goals

The Language Center's 1986 NEPT Committee Report (Al-Bader, et. al., 2) noted three main purposes of the new English-language testing system:

1. to diagnose relative English-language proficiency levels of all Kuwaiti secondary-school graduates
2. to help educators determine suitable admissions criteria for all colleges of Kuwait's various tertiary-level institutions, and
3. to appropriately place new students into English-language courses in their colleges.

Eventually, the report noted, the NEPT system was expected to "replace all current placement tests" (ibid., p.2) in use at the colleges of the various tertiary-level institutions in Kuwait.

C. Target Population and Test Content

Consistent with project goals, the target population, initially, for the NEPT project was the roughly 22,000 students graduating annually from Kuwait's Arabic-medium public and private secondary schools. Thus, test-form content in 1986-88 was based on the Crescent English series (Oxford University Press) taught between the fifth and

twelfth years of schooling in Kuwait.

For reasons of feasibility, the target population after 1988 became the approximately 3000-3500 new students (by law, at least 85 per cent Kuwaiti) admitted to Kuwait University each year. Test content thereafter was to be based primarily on the EFL curricula of the nine undergraduate colleges of the university, rather than on the content of the Crescent English series.

III. NEPT Specifications

A. NEPT Design Features

The NEPT Committee decided (Al-Bader, et. al., p.2) to include only three linguistic variables in the test system: grammatical structure, vocabulary and reading comprehension. Listening comprehension and essay writing parts were seriously considered for inclusion in the NEPT battery but were eventually rejected for reasons relating both to feasibility and practicality, on the one hand, and to professional measurement considerations, on the other.

The four-option multiple-choice (4MC) format was chosen for all three test parts, both because of years of highly successful use in the KU colleges' own English Placement Test (EPT) systems (see discussion in part B below) and because of recommendations of many professional experts (personal communications: Malcolm, 1978-80; Davies, 1986; and the NEPT project psychometric consultants Abou-Allam, Khammash and Ziomek). Each part was to contain 50 items,

all equally weighted at one point apiece. Timing of the more discrete-point items of parts 1 and 2, structure and vocabulary, was set at around 25 minutes each---i.e., at roughly two items per minute; whereas timing for the ten short reading passages of a test form, each followed by five related items on the same test page, was set at a very generous 80 minutes.

B. Professional Issues in NEPT Design

The Language Center had for years been using its 100-per cent 4MC EPTs in each college, with great success. These EPT forms had been professionally designed, written, reviewed and evaluated beginning in 1977 by Donald J. Malcolm (1977-80), former worldwide head of the TOEFL program at Educational Testing Services (ETS), in Princeton, N.J., in the U.S.A. All EPT forms and systems over the years had proven themselves highly reliable and regularly demonstrated their validity as placement instruments, in helping the colleges' EFL programs both to group new students effectively for language proficiency (Gaffney and Mason, 1983) and to predict their academic success (Mason 1984a and 1984b).

Decisions on content and format for the NEPT forms were made in full recognition of the strengths and the limitations of the 4MC format, of discrete-point testing, and of separate testing of the language skills and components in a test battery. The views of many measurement authorities

were taken into consideration. As to limitations of this more traditional type of EFL testing, see for example Hamp-Lyons, 1990, p. 298, in her review of Hughes, 1989; B.J. Carroll, 1980, p. 35; Clark, 1978, p. 24; Briere, 1972, p. 329; and Jones, 1979, p. 50. Yet, KULC staff over the years came to appreciate the strengths of well-designed and well-written 100-per cent 4MC testing systems like those of the TOEFL (Pike, 1979), of the International Language Testing Service (ILTS) of the British Council, and of their own EPT forms. The advantages of professionally developed objective testing systems have often been extolled by measurement authorities; see, for example, Stanley, 1965, p. 222; Gronlund, 1968, p. 26; Harris, 1969, p. 7; Valette, 1967, p. 173; Farhady, 1979; Ducroquet, 1980, pp. 242-3; Woodford, 1980, pp. 99-100; and Henning, 1982. And the advantages of separate-skills testing have been fully discussed by J.D. Carroll, 1972, p. 319.

The rapid, objective and highly reliable scoring of the EPT results at KULC, therefore, were seen not as the main benefits of the chosen testing format but simply as further evidence and support for the soundness, professionally speaking, of the decisions made concerning NEPT design. Further, the Language Center could readily demonstrate EPT and NEPT statistical accountability to KULC teaching staff,

to university officials and to visiting measurement experts, as regularly called for in the professional literature (Bachman, et. al, 1990; Alderson, 1990; Hamp-Lyons, 1990, p. 297).

Parts of the NEPT forms were all a suitable length (Stansfield, personal communication, 1990)---i.e., at least 40 or more items long---and proved themselves to be of around middle difficulty (Cooper, 1971), and were timed to permit all but the very weakest students to complete them (Malcolm, 1978-80). Without exception, the carefully designed EPT forms at KULC in the years 1978-90 achieved one of the most important goals for their teachers that a good English-language instructional program can strive for: to maximize the chances of teacher and student satisfaction and success (Gaffney and Mason, 1983, pp. 106-7)---and thus to minimize their frustration (Early and Bolitho, 1981, pp. 113-124; Ostojic, 1981, p. 46)---by creating reasonably homogeneous classes of new students all able to satisfy the stated learning objectives of their courses, learn together at roughly the same rate, and comprehend both their texts and the oral input of their entirely English-medium classroom.

C. Guidance to Staff for Quality Control

KULC staff assigned to write the NEPT items and to select and adapt suitable reading passages were provided many written guidelines as to 1) categories of structure, vocabulary and reading content to prepare and 2) profes-

sional references to turn to, to aid them in producing sound test material. The first set of guidelines was distributed to assure variety, validity and reliability of item and part content, while the second was required to attempt to assure staff conscientiousness in maintaining a high standard of quality of test output. Manuals recommended included those, for example, of Cohen, 1980; Harris, 1969; Heaton, 1975; Henning, 1987; B.J. Carroll, 1980; Malcolm, 1980; Oller, 1979; and Weir, 1988.

For the two discrete-point parts of structure and vocabulary, teachers were specifically asked to make their items as "communicative" as possible, with the excellent items of sections 2 and 3 of TOEFL forms as models of how sentence-level material can, with some thought, effort and care, be as meaningful and relevant to the real world as can passages of greater length. Teachers were reminded of Widdowson's discussion (1983, p. 101) on the importance of using teaching and testing material that contains "indexical" meaning; i.e., which has an instantly recognizable relationship and reference to the real world.

All test-content specifications and item-writing guidelines were issued by the KULC Tests and Measurement Office. That office also had responsibility for reviewing all new test material; modifying or rejecting it, as necessary; and assembling new NEPT forms according to project specifications. The office also wrote directives to all KULC staff

for the responsible and secure storage, administration, transport, handling and destruction of test forms, since loss of a single test booklet could be catastrophic to the NEPT system.

IV. Development of NEPT Forms

A. Item Production and Pretesting

In the 27 months between December 1986 and March 1989, KULC teachers created an item bank of nearly 2000 new NEPT items and 130 short reading passages (of between 75 and 150 words in length). Four major item-pretesting administrations were carried out in all Kuwait University EFL classes over a 24-month period, involving no less than 40 separate item pretest forms administered to a total of around 6000 students.

Superb cooperation and support were provided by the Kuwait University Computer Center staff for the rapid keypunching and analysis of all test results. Finally in 1990 keypunching was at last replaced by scoring of test results by optical mark reader (i.e., scanner) which dramatically reduced turnaround time for scoring and analysis from months to hours.

B. Steps in Test-Form Development

From the start the NEPT project consultants (Abou-Allam, Khamash, and later Ziomek) made clear that development of

each NEPT form could be expected to take no less than three to five years. Piloting of forms was to be followed by item validation, and by as many modifications and further administrations as necessary until item validity and subtest and test reliability had been established. Each stage required the involvement of random samples of a minimum number of representative student populations (at least 100 subjects for piloting and 1000 for each item validation administration).

Thereafter, form standardization would be carried out on around 5000 students (Khammash, personal communication, 1987). Each new form would be equated with already-standardized forms, and then would be ready for operational administration. With the demonstration that the NEPT results indicated end-of-semester predictive validity, the NEPT system would be declared operational.

Redefinition of the target student population in late 1988 led to the dissolution of the first three NEPT forms, the retirement of most items as too easy, and the writing of many new items to create more difficult new test forms. Piloting and item validation of these new forms began in early 1989. With the addition in 1989 of hundreds of items from the colleges' old EPT systems, six NEPT forms were under development by mid-1990.

V. Test Results

A. Fall 1987

The first full-length NEPT form (n=150) was administered on a trial basis in August 1987 to 1684 new students admitted to eight undergraduate colleges of Kuwait University. The form, whose items had been pretested the previous spring in all university EFL classes, was prepared as a base form (Form A) and two scrambled versions of the same items (Forms B and C), for in-class security reasons. Table 1 below presents means, standard deviations and reliabilities by college and form for total-test results---i.e., for part 1, 2 and 3 results combined.

As college-total means reveal, new students admitted to the science-based (and English-medium) colleges---Allied Health Sciences and Nursing; Engineering and Petroleum; Medicine; and Science---consistently outperformed new students enrolled in the Arabic-medium colleges teaching the arts, humanities and social sciences. That is, the students best prepared in English-language skills in the secondary schools were those most likely to be admitted to the university's more prestigious colleges teaching the theoretical and applied sciences. (Also, see Figures 1-2.)

Although this test was just a newly assembled form, its many strengths as a professionally developed assessment instrument can be seen in Tables 1-5. Total-university (N=1684) means for parts and test totals were very well-

balanced and almost exactly set at middle difficulty (Cooper, 1971, p. 341): 62.5 per cent. Test-total reliabilities for all colleges, separate and combined, were extremely high: .95-.97 (KR-20). Some 93 per cent of items were very effective in terms of difficulty level and discrimination power, and half were classified as highly effective (Table 2). Fully 95 per cent of item distractors worked well (see Table 3).

Further, NEPT inter-part correlations were all .84 or higher for all students combined (Table 4). And (Table 5) NEPT total-test results correlated with college EPT total-test scores at levels ranging from .80 to .95. (All college EPTs were 100-per cent 4MC forms except that of Allied Health, which employed a variety of more or less subjectively and objectively scored formats.)

B. 1988 Results in Secondary Schools

According to the original NEPT plan, KULC went forward with the initial steps in form development in early 1988 by administering the highly reliable first NEPT form to 629 fourth-year students in 26 Kuwaiti secondary schools. Data obtained from the 12 boys' and 14 girls' schools in Kuwait's five educational districts---Kuwait City Metropolitan, Hawalli, Ahmadi, Farwaniya, and Jahra Districts---produced notable differences among the randomly chosen schools, as well as among the randomly chosen classes in

the schools. Arabic-speaking EFL teachers from KULC carried out the administrations in the schools, to assure observance of proper testing and security procedures.

The results reported in this paper are probably unique in the Kuwait educational system for presenting valuable insights into EFL achievement at the secondary and tertiary levels of institutions supervised by Kuwait's educational authorities. Table 6 reveals that the 36 boys of the class at a private Hawalli secondary school (No. 13) scored at the level of the 203 male and female students admitted to Kuwait University's College of Engineering the previous fall on the same highly reliable test form (71.4 % vs. 71.1%). More impressively, this boys' class had a markedly lower standard deviation figure (22.66) than did the newly enrolled university group (28.76), suggesting that most of these high school boys had been learning English quite well, while many in the college group probably had not.

Of the eleven other boys' classes tested that spring, no other scored even at the 50 per cent level, and seven of them produced means below 40 per cent. However, of the fourteen randomly chosen girls' classes, nine of them scored above 50 per cent, and three of them, over 60 per cent. In short, the study provided convincing, if limited, support for what KULC teachers had known for years: among new university students, especially Kuwaiti students, girls almost always considerably outperformed boys from all educational districts of the country, and tended on

average to be much more motivated to work hard, to study and to learn the foreign language.

C. Subsequent NEPT Development

Subsequent EPT-NEPT administrations by KULC to new university students in the following two academic years yielded the same high standard of test results first seen in 1987: high part and form reliabilities, high percentages of effective items in all test parts and consistently high intra- and inter-test EPT-NEPT correlations (Tables 7a-c). With the shift in project orientation in late 1988, none of the NEPT forms was near completion, since a period of three to five years was needed to standardize each form, and new students entered the university only in the fall in numbers large enough to provide the population base needed to yield valid item, part and test results. However, the growing item bank of test questions and reading passages, plus the giant bank of item, part and test computer data continually increased the ability of KULC to "fine-tune" test items and parts, for parallel content, similar levels of difficulty, and therefore comparable effectiveness of corresponding parts, as well as for test totals.

VI. Conclusions

In short, the NEPT project was a kind of laboratory for both the KULC central administration and for its

teaching staff. In spite of substantial staff turnover each year, many teachers reported learning a great deal about issues of designing placement-test forms, defining valid test content, writing effective test items with the assistance of guidelines from handbooks written by internationally recognized testing authorities, and recognizing effective and ineffective test items, sections and parts.

Thus, in one sense a large, long-term testing project was undertaken, is now lost, and is therefore apparently no longer of importance to the international ELT community. In another sense, however, a large-scale cooperative educational enterprise was attempted, involving many thousands of students, hundreds of teachers and dozens of administrators and support staff in the university, the schools and the Ministry of Education. No effort was ever made to estimate the total number of man-hours---certainly in the tens of thousands---expended on the project over four years, much less of the dollar (or dinar) value that the project would have cost in another country without the resources to lavish on such a massive long-term undertaking.

All of the ELT professionals concerned benefited, directly or indirectly, and Kuwait's educational system certainly expected to benefit for a very large commitment of personnel and funding. The project is now past history, but it is worth remembering. May it, too, one day be revived.

References

Al-Bader, H., et. al. (1986). A Proposal for a General English Test for Secondary School Students. Internal report of Kuwait University Language Center.

Alderson, J.C. (1990). Comments as panel member on Cambridge-TOEFL Comparability Study at 1990 San Francisco TESOL Convention.

Bachman, L., Davidson, F. , & Foulkes, J. (1990). The Cambridge-TOEFL Comparability Study. Paper presented at 1990 San Francisco TESOL Convention.

Briere, E.J. (1972). Are we really measuring proficiency with our foreign language tests? In H. Allen & R.N. Campbell (eds.), Teaching English as a second language. New York: McGraw-Hill International Book Company.

Carroll, B.J. (1980). Testing communicative performance. Oxford: Pergamon Press.

Carroll, J.D. (1972). Fundamental Considerations in Testing for English Language Proficiency of Foreign Students. In H. Allen & R.N. Campbell (eds.), Teaching English as a second language. New York: McGraw-Hill International Book Company.

Clark, J.L.D. (1978). Psychometric Considerations in Language Testing: Advances in Language Testing, Series 2 (B. Spolsky, ed.). Arlington, Va.: Center for Applied Linguistics.

Cohen, A. (1980). Testing language ability in the classroom. Rowley, Mass.: Newbury House Publishers, Inc.

Cooper, R.L. (1971). Testing. In H. Allen & R.N. Campbell (eds.), Teaching English as a second language. New York: McGraw-Hill International Book Company.

Davies, A. (1986). Personal communication at 1986 Brighton, England, IATEFL conference that the 4MC format is "the best" EFL testing method for assessment of students' reading comprehension ability.

Ducroquet, L. (1980). Objective or Subjective Testing? System, 8(3), pp. 237-43.

Early, P., and Bolitho, R. (1981). "Reasons to be cheerful," or helping teachers to get problems into perspective.... System, 9(2), pp. 113-24.

Farhady, H. (1979). The Disjunctive Fallacy between Discrete-Point and Integrative Tests. TESOL Quarterly, 13(3), pp. 347-57.

Gaffney, J., & Mason, V.W. (1983). Rationalizing Placement and Promotion Decisions in a Major ELT Program. TESOL Quarterly, 17(1), 97-107.

Gronlund, N.E. (1968). Constructing achievement tests. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Hamp-Lyons, L. (1990). Recent Publications on Statistics, Language Testing, and Quantitative Research Methods II. TESOL Quarterly, 24(2), 293-300.

Harris, D.P. (1969). Testing English as a second language. New York: McGraw-Hill Book Company.

Heaton, J.B. (1975). Writing English language tests. London: Longman.

Henning, G. (1982). Twenty Common Testing Mistakes for EFL Teachers to Avoid. English Teaching Forum, 20(3).

Henning, G. (1987). A guide to language testing. Cambridge, Mass.: Newbury House Publishers.

Hughes, A. (1989). Testing for language teachers. Cambridge, Eng.: Cambridge University Press.

Jones, R.L. (1979). Performance Testing of Second Language Proficiency. In E.J. Briere & F.B. Hinofotis (eds.), Concepts in language testing. Washington, D.C.: Teachers of English to Speakers of Other Languages.

Malcolm, D.J. (1978-80). Many personal communications at Kuwait University Language Center on how best to design English placement test forms.

Malcolm, D.J. (1980). Test development manual. Kuwait: Kuwait University Press.

Mason, V.W. (1984a). Report on the Results of the English Placement Tests Administered in September 1983 in All Arabic-Medium Colleges. Internal KULC program report.

Mason, V.W. (1984b). Using Multiple-Choice Tests to Promote Homogeneity of Class Ability Levels in Large EGP and ESP Programs. System, 12(3), pp. 263-271.

Oller, J.W., Jr. (1979). Language Tests at school: a pragmatic approach. London: Longman.

Ostojic, B. (1981). Streaming in Foreign-Language Study. English Teaching Forum, 19(2), p. 46.

Pike, L.W. (1979). An evaluation of alternative item formats for testing English as a foreign language (TOEFL research report No. 2). Princeton, N.J.: Educational Testing Service.

Stansfield, C.W. (1990). Personal communication about the suitable length of valid 4MC parts of EFL test batteries.

Valette, R.M. (1967). Modern language testing. New York: Harcourt Brace Jovanovich, Inc.

Weir, C. (1988). Communicative language testing. Exeter: University of Exeter.

Widdowson, H.G. (1983). Competence and Capacity in Language Learning. On TESOL '82, pp. 97-106.

Woodford, P.E. (1980). Foreign Language Testing. The Modern Language Journal, 64(1).

Table 1: Total-Test Means, Standard Deviations and Reliabilities for All Colleges (n=150): NEPT Project, August 1987

College	Form & N	M (& %)	SD	Rel.
Allied Health	A: 41	88.8 (59.2)	29.97	.965
	B: 38	88.8 (59.2)	26.79	.959
	C: 26	94.1 (62.7)	28.33	.963
	T: 105	90.1 (60.1)	28.13	.961
Arts	A: 92	70.2 (46.8)	28.06	.957
	B: 91	70.6 (47.1)	30.74	.964
	C: 88	71.0 (47.3)	27.34	.955
	T: 271	70.6 (47.1)	28.76	.959
Commerce	A: 101	81.6 (54.4)	27.58	.957
	B: 108	83.0 (55.3)	30.46	.964
	C: 113	77.4 (51.6)	28.29	.958
	T: 322	80.7 (53.8)	29.02	.960
Engin-eering	A: 66	105.8 (70.5)	30.39	.970
	B: 66	108.2 (72.1)	28.21	.967
	C: 71	106.2 (70.8)	27.64	.965
	T: 203	106.7 (71.1)	28.76	.967
Law	A: 41	82.8 (55.2)	27.52	.957
	B: 43	83.8 (55.9)	28.58	.961
	C: 44	76.1 (50.7)	28.40	.959
	T: 128	80.9 (53.9)	28.16	.958
Medicine	A: 28	115.8 (77.2)	19.15	.942
	B: 29	119.0 (79.3)	20.97	.954
	C: 26	117.5 (78.3)	21.89	.958
	T: 83	117.4 (78.3)	20.34	.949
Science	A: 174	94.4 (62.9)	26.89	.958
	B: 176	92.9 (61.9)	30.12	.965
	C: 169	95.5 (63.7)	29.44	.965
	T: 519	94.3 (62.9)	28.87	.963
Shariah	A: 18	78.1 (52.1)	30.91	.966
	B: 17	90.8 (60.5)	27.85	.961
	C: 17	81.8 (54.5)	25.09	.951
	T: 52	83.5 (55.7)	27.95	.959
All Colleges Combined	A: 562	88.7 (59.1)	30.30	.965
	B: 568	89.5 (59.7)	31.81	.968
	C: 554	88.4 (58.9)	31.14	.967
	T: 1684	88.9 (59.3)	31.08	.966

Table 2: Effectiveness of Items, All Colleges Combined: NEPT Project, Fall 1987

Form & N	Part & n	No. (&%) of Effective Items: $\Delta \geq 8.5$ & $r \geq .40$	No. (&%) of Very Easy Items: $\Delta < 8.5$	No. (&%) of Non-Discriminating Items: $r < .35$	No. (&%) of Highly Effective Items: $\Delta \geq 11.0$ & $r \geq .50$
A: 562	1: 50	48 (96.0)	1 (2.0)	1 (2.0)	23 (46.0)
	2: 50	41 (82.0)	2 (4.0)	5 (10.0)	20 (40.0)
	3: 50	47 (94.0)	1 (2.0)	2 (4.0)	30 (60.0)
	T: 150	136 (90.7)	4 (2.7)	8 (5.3)	73 (48.7)
B: 568	1: 50	49 (98.0)	0	1 (2.0)	24 (48.0)
	2: 50	43 (86.0)	1 (2.0)	4 (8.0)	19 (38.0)
	3: 50	46 (92.0)	2 (4.0)	2 (4.0)	31 (62.0)
	T: 150	138 (92.0)	3 (2.0)	7 (4.7)	74 (49.3)
C: 554	1: 50	49 (98.0)	0	1 (2.0)	30 (60.0)
	2: 50	43 (86.0)	2 (4.0)	3 (6.0)	24 (48.0)
	3: 50	47 (94.0)	1 (2.0)	1 (2.0)	30 (60.0)
	T: 150	139 (92.7)	3 (2.0)	5 (3.3)	84 (56.0)
T: 1684	1: 50	49 (98.0)	0	1 (2.0)	27 (54.0)
	2: 50	43 (86.0)	2 (4.0)	5 (10.0)	20 (40.0)
	3: 50	47 (94.0)	1 (2.0)	2 (4.0)	29 (58.0)
	T: 150	139 (92.7)	3 (2.0)	8 (5.3)	76 (50.7)

Table 3: Effectiveness of Distractors, All Colleges Combined: NEPT Project, Fall 1987

Form & N	Part & n	No. of Distractors	No. (&%) of Weak or Non-Functioning Distractors *	No. (&%) of Functioning Distractors	No. (&%) of Items with Weak Distractors
A: 562	1: 50	150	15 (10.0)	135 (90.0)	14 (28.0)
	2: 50	150	9 (6.0)	141 (94.0)	6 (12.0)
	3: 50	150	7 (4.7)	143 (95.3)	6 (12.0)
	T: 150	450	31 (6.9)	419 (93.1)	26 (17.3)
B: 568	1: 50	150	11 (7.3)	139 (92.7)	10 (20.0)
	2: 50	150	6 (4.0)	144 (96.0)	4 (8.0)
	3: 50	150	8 (5.3)	142 (94.7)	6 (12.0)
	T: 150	450	25 (5.6)	425 (94.4)	20 (13.3)
C: 554	1: 50	150	12 (8.0)	138 (92.0)	10 (20.0)
	2: 50	150	10 (6.7)	140 (93.3)	7 (14.0)
	3: 50	150	11 (7.3)	139 (92.7)	8 (16.0)
	T: 150	450	33 (7.3)	417 (92.7)	25 (16.7)
Total: 1684	1: 50	150	11 (7.3)	139 (92.7)	11 (22.0)
	2: 50	150	6 (4.0)	144 (96.0)	4 (8.0)
	3: 50	150	7 (4.7)	143 (95.3)	5 (10.0)
	T: 150	450	24 (5.3)	426 (94.7)	20 (13.3)

*Defined here as those attracting fewer than 4.0 per cent of students.

Table 4: Correlation Coefficient Matrices for All Colleges Combined: NEPT Project, Fall 1987

Form & N	Variable	1	2	3	4
	Gr: 1	1.00			
A:	Vo: 2	.85	1.00		
562	RC: 3	.84	.84	1.00	
	To: 4	.95	.95	.95	1.00
	Gr: 1	1.00			
B:	Vo: 2	.87	1.00		
568	RC: 3	.85	.86	1.00	
	To: 4	.95	.95	.95	1.00
	Gr: 1	1.00			
C:	Vo: 2	.87	1.00		
554	RC: 3	.82	.83	1.00	
	To: 4	.95	.95	.93	1.00
	Gr: 1	1.00			
T:	Vo: 2	.86	1.00		
1684	RC: 3	.84	.84	1.00	
	To: 4	.95	.95	.94	1.00

Table 5: Intercorrelations of Part and Total Scores by College: NEPT Project, Fall 1987

College & N	Variable	obtained r	r corrected for attennation
Allied Health: 105	Grammar	--	--
	Vocabulary	--	--
	Reading	--	--
	Totals	.803	--
Arts: 271	Grammar	.848	.954
	Vocabulary	.809	.924
	Reading	.804	.913
	Totals	.901	.943
Com.: 322	Grammar	.903	(1.011)
	Vocabulary	.871	.995
	Reading	.859	.984
	Totals	.950	.994
Engin.: 203	Grammar	.800	.895
	Vocabulary	.766	.870
	Reading	.766	.871
	Totals	.859	(1.009)
Law: 128	Grammar	.862	.973
	Vocabulary	.818	.927
	Reading	.801	.913
	Totals	.904	.944
Med.: 83	Grammar	.731	.831
	Vocabulary	.736	.874
	Reading	.804	.887
	Totals	.872	.916
Sharia: 52	Grammar	.872	.986
	Vocabulary	.783	.888
	Reading	.742	.944
	Totals	.952	(1.012)

Table 6 : Total-Test Results for All 4th-Year Secondary-School Students: 8 February and 6 March 1988 Administrations (Schools 1-10 & 11-26, respectively)

District & School No.*	Gender & N	$\bar{M}$ (&%)	SD	Rel.
1: 1	B: 30	72.6 (48.4)	26.57	.953
1: 2	G: 31	75.1 (50.1)	20.86	.930
2: 3	B: 33	55.3 (36.9)	17.20	.890
2: 4	G: 30	90.4 (60.3)	24.17	.951
3: 5	B: 32	42.7 (28.5)	15.16	.874
3: 6	G: 35	87.5 (58.3)	21.98	.939
4: 7	B: 32	67.1 (44.7)	26.45	.952
4: 8	G: 33	72.2 (48.1)	24.73	.947
5: 9	B: 31	50.3 (33.5)	15.10	.865
5:10	G: 25	77.2 (51.5)	22.22	.939
0:11	G: 29	73.0 (48.7)	17.91	.903
0:12	G: 33	93.1 (62.1)	21.12	.938
0:13	B: 36	107.1 (71.4)	22.66	.950
0:14	B: 29	58.8 (39.2)	22.43	.934
1:15	G: 39	83.4 (55.6)	21.07	.931
1:16	B: 26	58.8 (39.2)	22.89	.938
1:17	G: 33	60.6 (40.4)	13.88	.839
1:18	G: 24	88.0 (58.7)	25.30	.953
2:19	B: 24	58.5 (39.0)	15.00	.820
2:20	G: 33	99.5 (66.3)	27.54	.964
2:21	G: 30	60.0 (40.0)	14.20	.846
3:22	G: 33	76.5 (51.0)	23.81	.943
3:23	B: 24	44.8 (29.9)	9.69	.689
3:24	B: 26	70.5 (47.0)	24.24	.945
5:25	B: 30	65.4 (43.6)	26.84	.954
5:26	G: 25	70.9 (47.3)	16.54	.888

*Schools designated with a zero for district are private schools.

Table 7a: Total Scores for NEPT Forms A + B (n=150), All Colleges: August 1988

College	N	$\bar{M}$ (& %)	SD	Rel.
A.H.	20	105.2 (70.1)	24.20	.958
Arts	25	72.6 (48.4)	27.62	.957
Com.	77	89.9 (59.9)	29.40	.964
Educ.	89	81.8 (54.5)	25.73	.951
Engin.	32	104.9 (69.9)	28.108	.968
Law	29	86.3 (57.5)	32.35	.969
Med.	15	112.6 (75.1)	31.00	.976
Sci.	65	99.5 (66.3)	28.32	.963
Total	352	91.3 (60.9)	29.88	.965

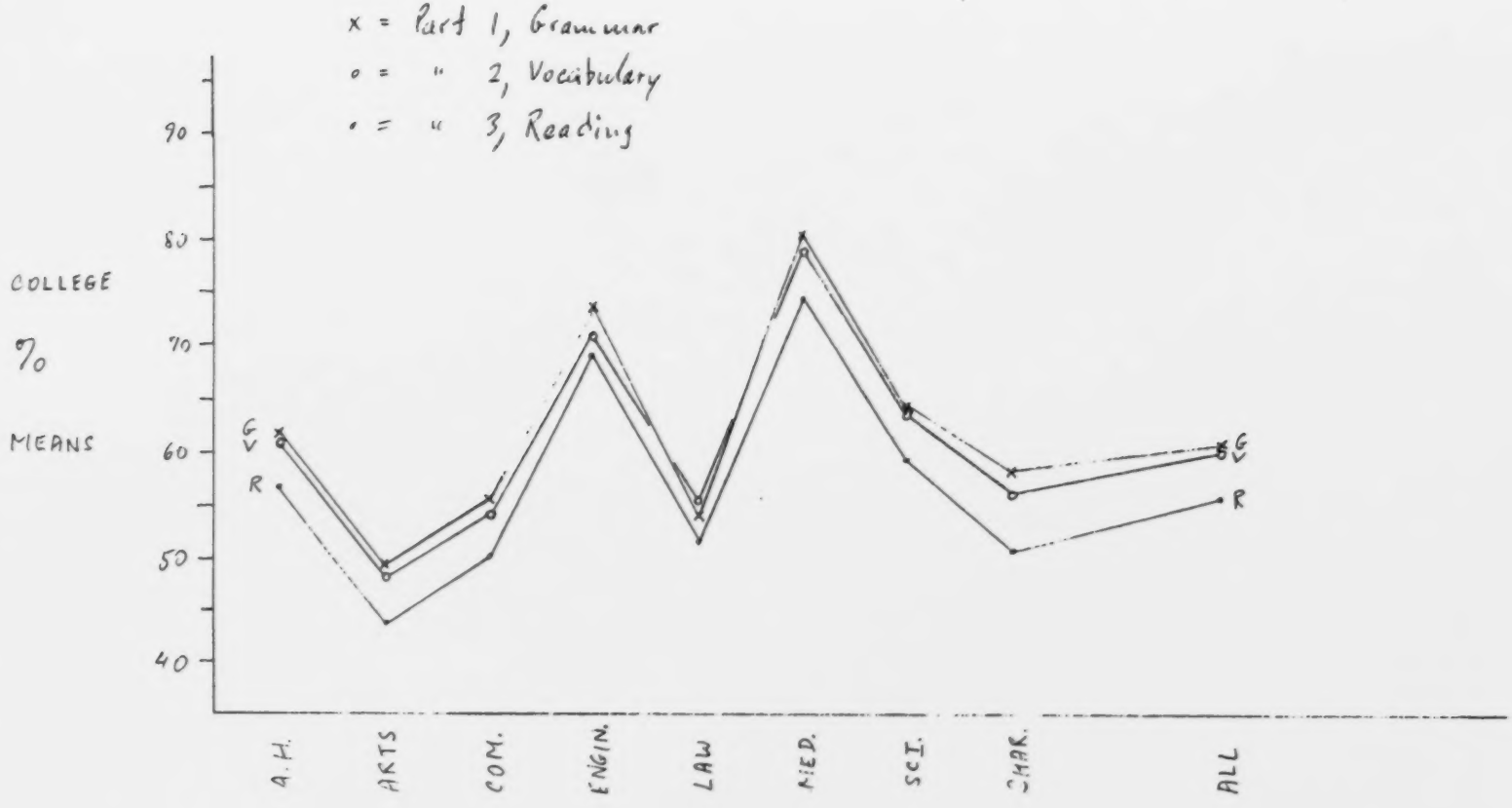
Table 7b: Total Scores for NEPT Form J (n=150), All Colleges: August 1988

College	N	$\bar{M}$ (& %)	SD	Rel.
A.H.	42	100.6 (67.1)	26.96	.963
Arts	48	78.6 (52.4)	27.62	.957
Com.	143	89.5 (59.7)	30.05	.965
Educ.	176	88.7 (59.1)	25.41	.952
Engin.	91	110.1 (73.4)	30.01	.972
Law	59	86.4 (57.6)	28.49	.960
Med.	29	115.2 (76.8)	25.44	.966
Sci.	126	105.5 (70.3)	26.40	.961
Total	714	95.5 (63.7)	29.48	.965

Table 7c: Total Scores for NEPT Form X (n=150), All Colleges: August 1988

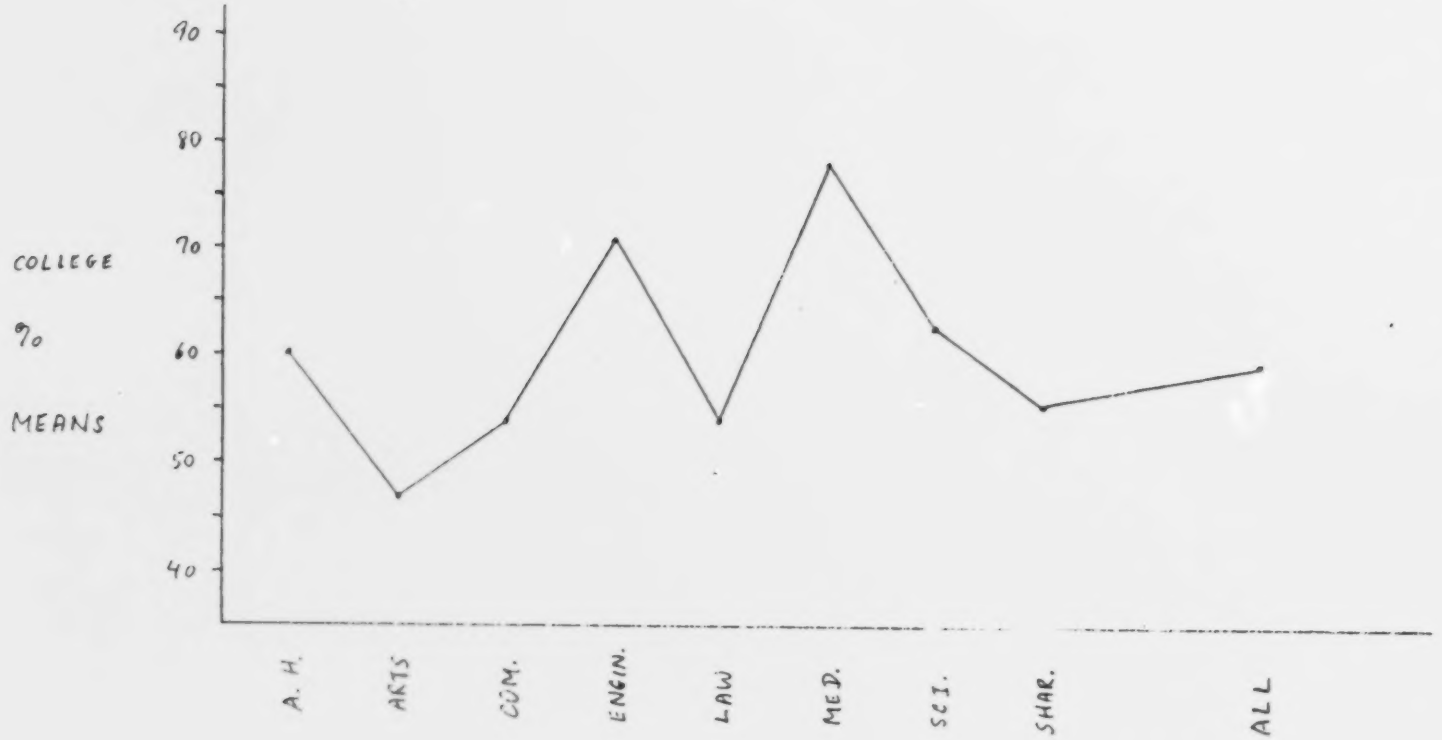
College	N	$\bar{M}$ (& %)	SD	Rel.
A.H.	41	98.8 (65.9)	33.12	.974
Arts	50	75.9 (50.6)	31.04	.965
Com.	147	87.0 (58.0)	29.50	.963
Educ.	178	82.3 (54.9)	26.23	.953
Engin.	81	107.0 (71.3)	27.81	.966
Law	59	84.2 (56.1)	29.46	.963
Med.	28	107.2 (71.5)	20.52	.941
Sci.	126	100.7 (67.1)	29.83	.968
Total	710	91.0 (60.7)	30.28	.966

FIG. 1 : COMBINED-LEVEL % MEANS ON PARTS 1-3, FOR ALL COLLEGES: NEPT PROJECT, FALL 1987



[NOV. 1987]

FIG. 2: COMBINED-LEVEL % MEANS FOR <u>TEST TOTALS</u> (n: 150)
FOR ALL COLLEGES: NEPT PROJECT, FALL 1987



[NOV. 1987]